DOCUMENT RESUME

ED 234 607 FL 013 919

AUTHOR Robertson, Daniel L.
TITLE Toward a Model for ESL Program Evaluation.
PUB DATE 82
NOTE 13p.; Paper presented at the Annual Convention of Teachers of English to Speakers of Other Languages (17th, Toronto, Ontario, Canada, March 15-20, 1983).
PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Education; *English (Second Language); Evaluation Methods; Higher Education; Models; *Program Evaluation; Second Language Programs

ABSTRACT

The variables that must be considered in English as a second language (ESL) program evaluation, major educational evaluation models, and a standards-based model for ESL program evaluation are discussed. Different ESL programs are examined including intensive programs and adult and university programs. Program variables such as subject matter, learner characteristics, academic setting, and length and intensity of training are addressed. Commonly used models for program evaluation are noted, such as systems analysis, behavioral objectives, management analysis, goal-free, art criticism, professional review, adversarial, and case study. The features of these models are related to problems presented by the ESL field. It is shown that, while all of these models can contribute to ESL evaluation, each is inadequate by itself to fairly evaluate ESL programs. A model for ESL program evaluation is presented which is based on standards for educational evaluations rather than on different kinds of evaluation approaches. This composite model draws upon features of other models and provides a mechanism for defining the evaluation problem and designing the evaluation. (Author/RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED234607

Toward a Model for ESL Program Evaluation

Daniel L. Robertson
University of Illinois

Abstract

This paper begins with a discussion of the variables which are important to the evaluation of a variety of ESL programs, such as intensive programs, adult programs, and university programs. The variables include subject matter, learner characteristics, academic setting, and length and intensity of training, as well as factors common to educational programs in general.

The second part of the paper briefly describes the most commonly-used models for educational evaluation, using the work of Stake (1974) and House (1980) as a guide. The models include systems analysis, behavioral objectives, management analysis, goal-free, art criticism, professional review, adversarial, and case study. The features of these models are related to problems presented by the field of ESL, and it is shown that while all of the models have something to offer the ESL evaluator, each is inadequate by itself for a fair and just evaluation of ESL programs.

The final part of the paper presents a model for ESL program evaluation which is based on standards for educational evaluations (Joint Committee on Standards for Educational Evaluation, 1981) rather than on different kinds of evaluation approaches. This shift of emphasis permits the use of methods from several different models, as long as they meet the requirements of the standards. The specific standards which are incorporated into the model are those which are most relevant to defining the evaluation problem and designing the evaluation.

FL013919

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Daniel Robertson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Toward a Model for ESL Program Evaluation*

Although ESL testing has long been a productive area of research in our field, there has been little published work in the area of program evaluation. The evaluation of ESL programs has, for the most part, fallen to ESL program administrators, who have also often been held responsible for the worth of their programs. The purpose of this paper is to enumerate some of the variables that must be considered in ESL program evaluation; briefly overview the major educational evaluation models, and propose a standards-based model for ESL program evaluation.

The Variables

The first variable to be considered is that of the subject matter. ESL courses often are categorized according to the use to which language is to be put once it is learned. Thus we have categories such as survival English, basic English, general English, conversational English, English for academic purposes, English for special purposes, English for science and technology, vocational English, technical English, business English and so on.

The second variable concerns the nature of the learners. Their age, of course, is very important. Their experience with English, either in their home country, another English speaking country, or the U.S. is an important factor. Their native languages, native countries, and cultural backgrounds must also be considered.

The academic setting of the learning is also very important. This setting may be elementary or secondary and include bilingual components. It may be a community college, college, or university. The settings for vocational or technical schools, commercial schools, business schools, or refugee centers may have quite different characteristics.

---

* Revised version of a paper which appeared in TESL Studies 5, 1982, and was presented at the TESOL '83 conference in Toronto.

The length and intensity of the training in English is another very important variable. Intensive, semi-intensive, supplemental, bilingual mainstream or maintenance programs would differ considerably in this regard.

There are many other variables which must be considered in the evaluation of ESL programs, of course. These are common to educational programs in general. They include budget, staff, curriculum, teaching materials and equipment, physical plant, quality of instructional program, research, location, climate, geography, and so on.

My purpose here is to note that ESL programs include variables which reflect their diversity as well as their similarity to other educational programs.

The Evaluation Models

In setting out to formalize the evaluation of programs which are so diverse, it is appropriate to consult the approaches which have been developed for the evaluation of educational programs in general, to see what features they have which might be appropriately included in the evaluation methodology.

Systems Analysis (Rivlin, 1971; Rossi, Freeman, and Wright, 1979) The systems analysis model is based on the idea that the way to find the truth is through scientific methodology. The output measures are limited in number and correlated with variables in program design. The model is designed for efficiency, and takes the perspective of the policy maker. Its weaknesses are that it assumes that one can assess the worth of a program with a few test scores from subjects in an educational experiment, that it fails to include the attitudes and feelings of the participants, and that it often relies on opaque statistics to enlighten the audiences.

This model would be useful in ESL for a limited number of programs. It would appeal to tightly-designed, limited-objective courses, where a satisfactory gain in scores on one or two measures would satisfy the requirements of the course. Since few ESL courses are so limited in scope, and since an ESL program represents such a variety of cultural backgrounds and involves so many interactions among diverse peoples, this model cannot alone satisfy the requirements of a fair evaluation of ESL programs. It is a very credible model, however, and the use of valid and reliable test scores lends obvious support to less objective instruments and measures in representing the outcomes of a program.

Behavioral Objectives (Tyler, 1950; Mager, 1962) The behavioral objectives model depends upon the precise specification of measurable goals and domain-referenced testing to demonstrate goal achievement. Like the systems analysis approach, this goal-based approach assumes that the methods of science can be applied to educational programs. These approaches fail to consider the interactions among people in the course of a program, as well as other effects which are not included in the specified goals.

This model would be particularly useful if it were incorporated into the ESL program design at the inception of the program. By carefully specifying all the goals of the program and designing tests and other measurement instruments to measure progress toward these goals, the evaluator could monitor the intended effects of the program more closely. The problem, of course, lies in the specification of the goals of the ESL program and the construction of valid tests of achievement. This type of evaluation component would be difficult to apply from the outside to an ongoing program, because it would require the

evaluator's work. The evaluator using this model must be coherent and persuasive, or his authority will be questioned, but he must be fair and just, or the usefulness and validity of his evaluation will be questioned. The audience for this type of evaluation must be considered, especially if it is large. A bad review of an educational program could be considered an indictment of the people in the program and could cause serious repercussions for all concerned.

This model could complement the more quantitative models with input from experts in the field of ESL. It suggests the need for experts in ESL to consider evaluation more seriously in order to illuminate some of the most important aspects of an ideal program for a given setting. The subjective nature of this model, however, would prevent its being used for the evaluation of most ESL programs, because the values of the participants would be likely to differ widely from those of the evaluator.

Professional Review (National Study of School Evaluation, 1978; Evaluative Criteria) The professional review model for educational evaluation depends upon professionals for the criteria and standards to be applied. It provides a holistic assessment of programs, usually by first having the program staff complete a self-evaluation checklist, then following up with a short visit by a team of professionals. This type of evaluation can be useful in solving crisis situations or preserving institutions, and also results in self-awareness among program staff as a result of self-evaluation. There are also problems, however. It may be that professionals are less objective than they should be when evaluating other professionals. Also, great differences often exist in the criteria used and their weightings, the visits of accreditation teams vary considerably, and maintaining confidentiality is difficult. This model can be quite effective if the procedures and processes of the evaluation effort are clearly specified at the outset.

Because this model provides the most versatile approach to widely divergent programs, it can be of great use in evaluating ESL programs. In order to accurately judge a program, one must understand it, and this model provides an effective way to reach that understanding. Despite the qualitative and subjective nature of this model and the drawbacks I have mentioned, it has much to offer the ESL evaluator. When used in conjunction with other, more quantitative approaches, it can provide a meaningful backdrop against which numbers may be made more understandable.

It may be seen from this discussion that all of the models have something to offer the ESL evaluator, yet each is inadequate by itself for the evaluation of ESL programs. Every program is unique, and a fair evaluation of that program will require an evaluation model which is responsive to its characteristics. The next part of this paper will suggest a model for evaluation of ESL programs which is based on standards for evaluations, rather than different kinds of evaluation approaches. This shift of emphasis will permit the use of the methods from several different models, as long as they meet the requirements of the standards.

A Standards-Based Model

This model is based on standards which have been developed for the evaluation of educational programs, projects, and materials. (Joint Committee on Standards for Educational Evaluation, 1981). The standards fall into four general categories: utility, feasibility, propriety, and accuracy. The specific standards which I have incorporated into this model are those which are most relevant to defining the evaluation problem and designing the evaluation. In the process of evaluating, of course, many of the other standards will become applicable to any evaluation effort. My purpose here is to show how the standards may be used as a model for the evaluation of an ESL program, and

provide the prospective ESL evaluator with a basic guide to evaluating an ESL program.

Audience identification is the essential first step in ESL program evaluation. The audiences for any given evaluation may differ, of course. The audiences may include the learners themselves, the ESL program staff, the school staff, language educators, minority groups, and the public at large. The identification of these groups is absolutely essential in determining the scope and focus of the evaluation.

The program to be evaluated must be described in detail. The description of the program should include all of those areas which have been mentioned above as variables in ESL programs: (1) the area of ESL and the subject matter of the program should be clearly described; (2) the characteristics of the learners, including age, experience with English, home country, native language, and cultural background should be investigated; (3) the academic setting of the program, whether elementary, secondary, post-secondary, or adult academic, technical, or vocational, should be clearly and completely described; (4) the length and intensity of the program should be determined and described in detail; and (5) other characteristics of the program, such as staff qualifications, physical setting, and curricular goals, should be described.

The context in which the program exists needs to be clearly set forth. This would include a description of the social, political, economic, and linguistic aspects of the environment, as well as a determination of how well the program fits with its environment.

Once the audiences, the program, and the context have been clearly identified and described, questions must be formulated to focus the evaluation. It is essential that these questions be responsive to the needs of the audiences.

It is important in most cases to specify in writing the contract which exists between client and evaluator, not only for the protection of those involved, but also for the understanding of all parties of the plan for the evaluation. In the formal contract, a number of different but related areas need to be examined and described. The objectives of the evaluation and the questions to be investigated need to be clearly stated. The procedures for data collection and analysis should be specified. The reporting plan and bias control measures need to be clearly described. Contributions of the client, in both supplies and personnel support, should be mentioned. Guidelines for the plan of work, as well as for the amendment or termination of the contract, should be clearly stated. The contract should include a budget for the evaluation. It should also be examined for accordance to local, state, and federal laws. Finally, after negotiation and agreement on its contents, it should be signed and copies retained by the parties involved.

The validity of the information on which the evaluation is to be based must be clearly established. The instruments and procedures should be checked against the objectives and content of the program. Judgments regarding their validity should be obtained from both participants in the program and outside subject-matter experts. The reasons for selection of specific instruments and procedures should be detailed, and the validity of all the instruments and procedures should be established vis-a-vis the questions addressed in the evaluation. Special attention should be given to new measurement instruments, and possible misinterpretation of measures or scores should be pointed out.

The instruments chosen for data collection should have acceptable reliability for the uses to which they are put. Methods of estimating reliability should be appropriate and defensible. The effects of the setting and the

sample on the reliability of the information should be recognized, and the measurement techniques should be clearly described so that the audiences may make their own judgments regarding reliability.

The quantitative information which is collected must be analyzed in order to support the interpretations to be made. The analyses must be systematic, and should proceed in this order: organize, summarize, interpret, report. Independent sets of data should be collected and analyzed, and potential weaknesses in the collection or analysis of the data should be reported.

The qualitative information also must be analyzed in order to support the interpretations to be made. Both the analysis procedure and the method of summarization are important in this regard. Confirmatory evidence must be sought. Not only should different types of information be gathered, but the categories of information should be meaningful, internally consistent, and mutually exclusive. Collection of qualitative information should be limited when sources are exhausted or when extensive regularity or redundancy of information is encountered. Potential weaknesses in the collection or analysis of the data should be checked with the audiences of the evaluation.

The conclusions reached in an evaluation should be both defensible and defended in the final evaluation report. The conclusions should be based on sound logic and appropriate information. The conclusions should be defended by an accounting of the procedures, information, and assumptions of the evaluator. Possible alternative explanations should be discussed, as well as the reasons for their rejection. Information which existed prior to the evaluation should be used for support. The conclusions should be related to the questions of the audiences, and the audiences should be advised on the interpretation of equivocal findings.

Because of the specification of the bases for value judgments, the specification of purposes and procedures, and the formalization of the evaluation contract, the findings and reports should have adequate safeguards against bias. It is important, however, to seek out and report possible sources of bias, as well as conflicting points of view regarding the conclusions and recommendations. It is also important for the evaluator to establish and maintain his independence throughout the evaluation effort.

Each of these steps provides an essential ingredient for the fair evaluation of programs in ESL. Because of the nature of such programs, the first three steps are particularly important. The other steps are as necessary for ESL program evaluations as they are for any other educational evaluations. By following these guidelines, it may be possible for the evaluator to evaluate ESL programs effectively and fairly.

REFERENCES

Eisner, E. 1979. The Educational Imagination. New York: Macmillan.

Guba, E. G. 1978. Toward a Methodology of Naturalistic Inquiry in Educational Evaluation. Los Angeles: Center for the Study of Evaluation, UCLA.

Guttentag, M. 1973. Subjectivity and its use in evaluation research. Evaluation 1, 2: 60-65.

House, Ernest R. 1980. Evaluating with Validity. Beverly Hills: Sage.

Joint Committee on Standards for Educational Evaluation 1981. Standards for Evaluation of Educational Programs, Projects, and Materials. New York: McGraw-Hill.

Kreidler, Carol J. (Ed.) 1981. Reports of Ad-Hoc Committee on Employment Issues. (Mimeo, March 4, 1981), Teachers of English to Speakers of Other Languages.

Mager, R. F. 1962. Preparing Objectives for Programmed Instruction. San Francisco: Fearon.

--- 1972. Goal Analysis. Belmont, CA: Fearon.

National Study of School Evaluation 1978. Evaluative Criteria. Arlington, VA: Author.

National Study of Secondary School Evaluation 1969. Evaluative Criteria. Washington, DC: Author.

Owens, T. R. 1973. Educational evaluation by adversary proceeding. In E. R. House (Ed.) School Evaluation. Berkeley: McCutchan.

Patton, M. Q. 1979. Utilization-Focused Evaluation. Beverly Hills: Sage.

Rivlin, A. M. 1971. Systematic Thinking for Social Action. Washington, DC: Brookings Institution.

Rossi, P. H., H. E. Freeman, and S. R. Wright 1979. Evaluation: A Systematic Approach. Beverly Hills: Sage.

Scriven, M. 1973. Goal free evaluation. In E. R. House (Ed.) School Evaluation. Berkeley: McCutchan.

Stake, R. E. 1974. Nine Approaches to Educational Evaluation. Urbana: University of Illinois Center for Instructional Research and Curriculum Evaluation (mimeo).

--- 1975. Evaluating the Arts in Education: A Responsive Approach. Columbus, OH: Merrill.

--- 1978. The case study method in social inquiry. Educational Researcher 7:5-8.

Stufflebeam, D. L. 1969. Evaluation as enlightenment for decision-making. In Beatty (Ed.) Improving Educational Assessment and an Inventory of Affective Behaviors. Washington, DC: Association for Supervision and Curriculum Development, NEA.

--- 1973. An introduction to the PDK book. In B. Worthen and J. R. Sanders (Eds.) Educational Evaluation: Theory and Practice, Worthington, OH: Charles A. Jones.

Tyler, R. W. 1950. Basic Principles of Curriculum and Instruction. Chicago: Univ. of Chicago Press.

Wolf, R. L. 1975. Trial by jury: The process. Phi Delta Kappan 57:185-187.